

(19)



(11)

EP 4 075 327 A1

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:

19.10.2022 Bulletin 2022/42

(51) International Patent Classification (IPC):

G06K 9/62 ^(2022.01)

G06N 3/08 ^(2006.01)

G06N 3/04 ^(2006.01)

(21) Application number: **21168669.6**

(52) Cooperative Patent Classification (CPC):

G06N 3/08; G06K 9/6256; G06K 9/6271;

G06K 9/6288; G06N 3/0454; G06N 3/0481

(22) Date of filing: **15.04.2021**

(84) Designated Contracting States:

**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB
GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO
PL PT RO RS SE SI SK SM TR**

Designated Extension States:

BA ME

Designated Validation States:

KH MA MD TN

(72) Inventors:

- **LELOWICZ, Kamil**
44-295 Bogunice (PL)
- **NOWAK, Mariusz**
40-145 Katowice (PL)

(74) Representative: **Manitz Finsterwald**

Patent- und Rechtsanwaltspartnerschaft mbB
Martin-Greif-Strasse 1
80336 München (DE)

(71) Applicant: **Aptiv Technologies Limited**
St. Michael (BB)

(54) **METHODS AND SYSTEMS FOR TRAINING A MACHINE LEARNING METHOD FOR DETERMINING PRE-DETERMINED POINTS IN AN IMAGE**

(57) A computer implemented method for training a machine learning method for determining pre-determined points in an image comprises the following steps carried out by computer hardware components: training a first machine learning method based on training data indicating whether or not a pre-determined point is present in a cell of the image; training a second machine

learning method based on a position of pre-determined points in the cell and based on training of the first machine learning method; and training a third machine learning method based on a modified version of the image and based on training of the second machine learning method.

EP 4 075 327 A1

Description

FIELD

[0001] The present disclosure relates to methods and systems for training a machine learning method for determining pre-determined points in an image.

BACKGROUND

[0002] Digital imaging devices, such as digital cameras, are used in automotive applications for a variety of tasks. However, it may be required to properly calibrate the digital imaging device.

[0003] Accordingly, there is a need to provide efficient and reliable methods for camera calibration.

SUMMARY

[0004] The present disclosure provides a computer implemented method, a computer system and a non-transitory computer readable medium according to the independent claims. Embodiments are given in the sub-claims, the description and the drawings.

[0005] In one aspect, the present disclosure is directed at a computer implemented method for training a (third) machine learning method for determining pre-determined points in an image, the method comprising the following steps performed (in other words: carried out) by computer hardware components: training a first machine learning method based on training data indicating whether or not a pre-determined point is present in a cell of the image; training a second machine learning method based on a position of pre-determined points in the cell and based on training of the first machine learning method; and training a third machine learning method based on a modified version of the image and based on training of the second machine learning method.

[0006] In other words, overall training of a machine learning method may be split into three phases, wherein in each phase an individual machine learning method is trained, based on the training of the machine learning method of the previous phase (it will be understood that in the first phase no previous phase is available, and as such, the training in the first phase does not involve any previously trained neural network). The complexity of the training may increase from phase to phase.

[0007] For example, one or more of the first machine learning method, second machine learning method, and third machine learning method may be the machine learning method as described herein.

[0008] According to an embodiment, the modified version comprises image after image processing. The modified version of the image may also be referred to as an adversarial-like image. The adversarial-like image may be created to intentionally confuse the method, and to provoke wrong results.

[0009] According to an embodiment, the image

processing comprises at least one of a rotation, a scaling, a flipping, and an embedding in another image (for example in a random image).

[0010] According to an embodiment, the image comprises (or shows) a checkerboard.

[0011] According to an embodiment, the pre-determined points comprise saddle points of the checkerboard.

[0012] According to an embodiment, saddle points are crossings of black-white borders of the checkerboard.

[0013] According to an embodiment, the machine learning method comprises an artificial neural network.

[0014] In another aspect, the present disclosure is directed at a computer implemented method for determining pre-determined points in an image, the method comprising the following steps carried out by computer hardware components: providing the image to a sequence of processing layers; and determining the pre-determined points in the image based on the output of the last processing layer of the sequence of processing layers; wherein the last processing layer comprises a respective output data set for each of a plurality of cells of the image; wherein each output data set comprises a plurality of probabilities and a plurality of positions, each probability indicating a probability of a pre-determined number of pre-determined points in the respective cell, and each position indicating a position in the respective cell.

[0015] In other words, a machine learning method may be provided which outputs respective probabilities for how many pre-determined points are present, and estimations of the position(s) of the pre-determined point(s).

[0016] The respective position may be a position of a potential pre-determined point.

[0017] The plurality of layers may include or may be 15 layers (which may form an artificial neural network or a part of an artificial neural network).

[0018] According to an embodiment, the output is of a YOLO type.

[0019] The output of a YOLO type may be a grid. Each cell in the grid may include information about detection. Each cell may include one or more detections. According to various embodiments, a single detection may be described by three parameters: (1) probability of existence of a saddle point in a particular cell, (2) saddle point position in the x axis, and (3) saddle point position in the y axis within the cell (range from 0 to 1, wherein (0,0) may mean top left cell corner and (1,1) may mean bottom right cell corner).

[0020] In other words, YOLO like output may include dividing the image into regions which are described by an output grid.

[0021] According to an embodiment, the last processing layer comprises a softmax layer.

[0022] According to an embodiment, the plurality of probabilities comprises an integer number of probabilities, wherein the pre-determined numbers of pre-determined points in the respective cell comprises integer numbers between zero and one less than the integer

number of probabilities.

[0023] According to an embodiment, the image comprises (or shows) a checkerboard.

[0024] According to an embodiment, the pre-determined points comprise saddle points of the checkerboard.

[0025] According to an embodiment, saddle points are crossings of black-white borders of the checkerboard.

[0026] In another aspect, the present disclosure is directed at a computer system, said computer system comprising a plurality of computer hardware components configured to carry out several or all steps of the computer implemented method described herein.

[0027] The computer system may comprise a plurality of computer hardware components (for example a processor, for example processing unit or processing network, at least one memory, for example memory unit or memory network, and at least one non-transitory data storage). It will be understood that further computer hardware components may be provided and used for carrying out steps of the computer implemented method in the computer system. The non-transitory data storage and/or the memory unit may comprise a computer program for instructing the computer to perform several or all steps or aspects of one or more of the computer implemented methods described herein, for example using the processing unit and the at least one memory unit.

[0028] According to an embodiment, the computer system further comprises a camera configured to acquire the image.

[0029] In another aspect, the present disclosure is directed at a non-transitory computer readable medium comprising instructions for carrying out several or all steps or aspects of one or more of the computer implemented methods described herein. The computer readable medium may be configured as: an optical medium, such as a compact disc (CD) or a digital versatile disk (DVD); a magnetic medium, such as a hard disk drive (HDD); a solid state drive (SSD); a read only memory (ROM), such as a flash memory; or the like. Furthermore, the computer readable medium may be configured as a data storage that is accessible via a data connection, such as an internet connection. The computer readable medium may, for example, be an online data repository or a cloud storage.

[0030] The present disclosure is also directed at a computer program for instructing a computer to perform several or all steps or aspects of one or more of the computer implemented methods described herein.

[0031] With the methods and devices as described herein, a technique of using machine learning to identify saddle points in camera calibration patterns may be provided.

DRAWINGS

[0032] Exemplary embodiments and functions of the present disclosure are described herein in conjunction

with the following drawings, showing schematically:

- Fig. 1A an illustration of an example of input data;
- 5 Fig. 1 B an illustration of a further example of input data;
- Fig. 2 an illustration of saddle point examples;
- 10 Fig. 3 an illustration of an output structure according to various embodiments;
- Fig. 4 an illustration of an adversarial like data example;
- 15 Fig. 5 a flow diagram illustrating a method for training a machine learning method for determining pre-determined points in an image according to various embodiments; and
- 20 Fig. 6 shows a flow diagram illustrating a method for determining pre-determined points in an image according to various embodiments.

25 DETAILED DESCRIPTION

[0033] According to various embodiments, a machine learning based saddle points detector may be provided.

[0034] Saddle point on an image (or picture) may be one of the most characteristic features that can be extracted from the image. Saddle points may be widely used for intrinsic calibration of a camera, extrinsic calibration of a camera, and/or matching characteristic points.

[0035] An accuracy of finding saddle points may have critical impact to quality of camera calibration. Moreover, each camera mounted in a car may need to be calibrated. The main goal of a method may be to find all saddle points in a picture with subpixel accuracy under all light conditions.

[0036] According to various embodiments, the input may be an image with one grayscale channel. The maximum resolution may be smaller or equal to 1920x1080; otherwise, the input image may be divided in smaller part. Each pixel may be coded as a float32.

[0037] Fig. 1A shows an illustration 100 of an example of input data, i.e. an input image. It can be seen how big distortion appears in the picture. The image shows a checkerboard.

[0038] Fig. 1B shows an illustration 150 of a further example of input data. It can be seen that checkerboards may be provided in more than one plane.

[0039] Fig. 2 shows an illustration 200 of saddle point examples. In the top portion of Fig. 2, the pixel values are shown over the pixel coordinates. The bottom portion shows the image, and the saddle point is at the crossing between the borders between dark areas and light areas.

[0040] According to various embodiments, a machine learning based saddle point detector may be provided

and may improve robustness of algorithm under all light conditions and may reduce execution time.

[0041] According to various embodiments, the (artificial) neural network may be provided.

[0042] According to various embodiments, a training method of an (artificial) neural network may be provided, and a training stage may utilize adversarial-like samples, as will be described in more detail below.

[0043] According to various embodiments, the core of neural network may be based on residual neural network (ResNet). For example, the network may include (or may be made up) of 15 layers and may contain about 1 million parameters. The input to the neural network may be a grayscale picture with shape (or dimensions) (1080, 1920); in other words, the input image may be grayscale image which is 1920 pixels wide and 1080 pixels high. The output from the neural network may be YOLO like (YOU LOOK ONLY ONCE). The grid may have size 135 by 240 (64 times smaller than the input). Each cell may consist of 7 numbers, that describe corresponding 8px by 8px on original image. First three numbers may describe probability of no saddle point in cell, one saddle point in cell or 2 saddle points in a cell. These values are calculated using a softmax layer. The next 4 numbers may describe position of saddle points (x, y) for first and second detection. The position may be measured from upper left corner of the cell. The x and y may be in the range from 0 to 1, they are outputted using hard-sigmoid function.

[0044] Fig. 3 shows an illustration 300 of an output structure according to various embodiments. The input image may be divided into a plurality of cells, one of which is denoted exemplarily by reference sign 302. Each cell may be 8 pixels wide and 8 pixels high. An example of a saddle point 304 is illustrated.

[0045] According to various embodiments, the training of the neural network may include three phases:

1. In a first phase, only the confidence whether the saddle point/points exist in a cell or not may be taken into account.
2. In a second phase, the saddle point position may be included during training.
3. In a third (in other words: last) stage, in order to minimize the number of false positive cases, adversarial-like training may be used. The random parts of checkerboards may be rotated, scaled, flipped and then embedded in random pictures.

[0046] All three phases may be performed on the same neural network (i.e. each phase may modify the preliminary weights obtained from the previous phase)

[0047] In the first phase, the core of the neural network may be trained to only correctly predict existence of saddle point in the cells. When the network performs well in predicting the existence of saddle points, the resulting weights may be used as a starting point for the second phase.

[0048] In the second stage, the NN may also be trained to predict the position of each detected saddle point.

[0049] The third phase, may be the adversarial training phase, where the network may be fed with input that consists of parts of the checkerboard and real life images stitched together.

[0050] Fig. 4 shows an illustration 400 of an adversarial like data example.

[0051] In good light conditions and when the checkerboard is placed in the center of the image (and hence the distortion is small) and is perpendicular to the camera, the method according to various embodiments may find all or almost all the saddle points. Even if for example the floor, especially with a carpet, may be noisy, no false positive detections may be observed. Thus, under favorable conditions, the method according to various embodiments works very well.

[0052] Even situations where the checkerboard is no longer perpendicular to camera may not affect the performance of the method according to various embodiments.

[0053] Even when the checkerboard is off-center (from the camera central axis), as there are distorted corners, the method according to various embodiments may still perform very good and all or almost all the points may be detected.

[0054] Even when the lighting conditions are harder, and very bright background (for example behind a window) and dark interior lead to smaller contrast on checkerboard, so that the checkerboard lighting is not uniform, the method according to various embodiments may still perform very well.

[0055] The method according to various embodiments may be fast and may give good numerical results.

[0056] The method according to various embodiments may be very accurate in corners where distortion is very high and contrast very low.

[0057] The method according to various embodiments may be very robust.

[0058] Fig. 5 shows a flow diagram 500 illustrating a method for training a machine learning method for determining pre-determined points in an image according to various embodiments. At 502, a first machine learning method may be trained based on training data indicating whether or not a pre-determined point is present in a cell of the image. At 504, a second machine learning method may be trained based on a position of pre-determined points in the cell and based on training of the first machine learning method. At 506, a third machine learning method may be trained based on a modified version of the image and based on training of the second machine learning method.

[0059] According to various embodiments, the modified version may include or may be the input image after image processing.

[0060] According to various embodiments, the image processing may include or may be at least one of a rotation, a scaling, a flipping, and an embedding in another

image

[0061] According to various embodiments, the image may include or may show a checkerboard.

[0062] According to various embodiments, the pre-determined points may include or may be saddle points of the checkerboard.

[0063] According to various embodiments, saddle points may be crossings of black-white borders of the checkerboard.

[0064] According to various embodiments, the machine learning method may include or may be an artificial neural network.

[0065] Fig. 6 shows a flow diagram 600 illustrating a method for determining pre-determined points in an image according to various embodiments. At 602, the image may be provided to a sequence of processing layers. At 604, the pre-determined points in the image may be determined based on the output of the last processing layer of the sequence of processing layers. The last processing layer may include or may be a respective output data set for each of a plurality of cells of the image. Each output data set may include or may be a plurality of probabilities and a plurality of positions, each probability indicating a probability of a pre-determined number of pre-determined points in the respective cell, and each position indicating a position in the respective cell.

[0066] According to various embodiments, the output may be of a YOLO type.

[0067] According to various embodiments, the last processing layer may include or may be a softmax layer.

[0068] According to various embodiments, the plurality of probabilities may include or may be an integer number of probabilities, wherein the pre-determined numbers of pre-determined points in the respective cell comprises integer numbers between zero and one less than the integer number of probabilities.

[0069] According to various embodiments, the image may include or may show a checkerboard.

[0070] Each of the steps 502, 504, 506, 602, 604 and the further steps described above may be performed by computer hardware components.

Reference numeral list

[0071]

100 illustration of an example of input data

150 illustration of a further example of input data

200 illustration of saddle point examples

300 illustration of an output structure according to various embodiments

302 cell

304 saddle point

400 illustration of an adversarial like data example

500 flow diagram illustrating a method for training a machine learning method for determining pre-determined points in an image according to various embodiments

5 502 step of training a first machine learning method based on training data indicating whether or not a pre-determined point is present in a cell of the image

10 504 step of training a second machine learning method based on a position of pre-determined points in the cell and based on training of the first machine learning method

506 step of training a third machine learning method based on a modified version of the image and based on training of the second machine learning method

600 flow diagram illustrating a method for determining pre-determined points in an image according to various embodiments

20 602 step of providing the image to a sequence of processing layers

25 604 determining the pre-determined points in the image based on the output of the last processing layer of the sequence of processing layers

Claims

1. Computer implemented method for training a machine learning method for determining pre-determined points in an image, the method comprising the following steps carried out by computer hardware components:

- training (502) a first machine learning method based on training data indicating whether or not a pre-determined point is present in a cell of the image;

- training (504) a second machine learning method based on a position of pre-determined points in the cell and based on training of the first machine learning method; and

- training (506) a third machine learning method based on a modified version of the image and based on training of the second machine learning method.

2. The computer implemented method of claim 1, wherein the modified version comprises image after image processing.

3. The computer implemented method of claim 2, wherein the image processing comprises at least one of a rotation, a scaling, a flipping, and an embedding in another image

4. The computer implemented method of at least one

of claims 1 to 3,
wherein the image comprises a checkerboard.

5. The computer implemented method of claim 4,
wherein the pre-determined points comprise saddle
points of the checkerboard. 5
6. The computer implemented method of claim 5,
wherein saddle points are crossings of black-white
borders of the checkerboard. 10
7. The computer implemented method of at least one
of claims 1 to 6,
wherein the machine learning method comprises an
artificial neural network. 15
8. Computer implemented method for determining pre-
determined points in an image,
the method comprising the following steps carried
out by computer hardware components: 20
 - providing (602) the image to a sequence of
processing layers; and
 - determining (604) the pre-determined points in
the image based on the output of the last
processing layer of the sequence of processing
layers; 25wherein the last processing layer comprises a re-
spective output data set for each of a plurality of cells
of the image; 30
wherein each output data set comprises a plurality
of probabilities and a plurality of positions, each prob-
ability indicating a probability of a pre-determined
number of pre-determined points in the respective
cell, and each position indicating a position in the
respective cell. 35
9. The computer implemented method of claim 8,
wherein the output is of a YOLO type. 40
10. The computer implemented method of at least one
of claims 8 to 9,
wherein the last processing layer comprises a soft-
max layer. 45
11. The computer implemented method of at least one
of claims 8 to 10,
wherein the plurality of probabilities comprises an
integer number of probabilities, wherein the pre-de-
termined numbers of pre-determined points in the
respective cell comprises integer numbers between
zero and one less than the integer number of prob-
abilities. 50
12. The computer implemented method of at least one
of claims 1 to 11,
wherein the image comprises a checkerboard. 55

13. Computer system, the computer system comprising
a plurality of computer hardware components con-
figured to carry out steps of the computer implement-
ed method of at least one of claims 1 to 12.

14. The computer system of claim 13, further compris-
ing:
a camera configured to acquire the image.

15. Non-transitory computer readable medium compris-
ing instructions for carrying out the computer imple-
mented method of at least one of claims 1 to 12.

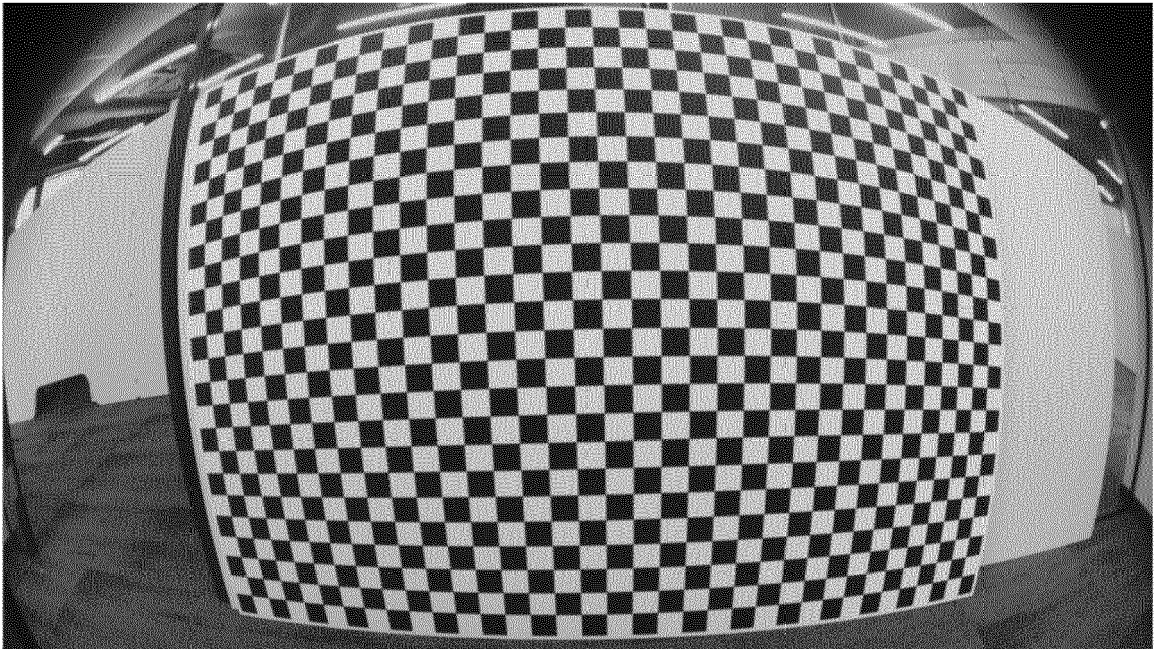


Fig. 1A

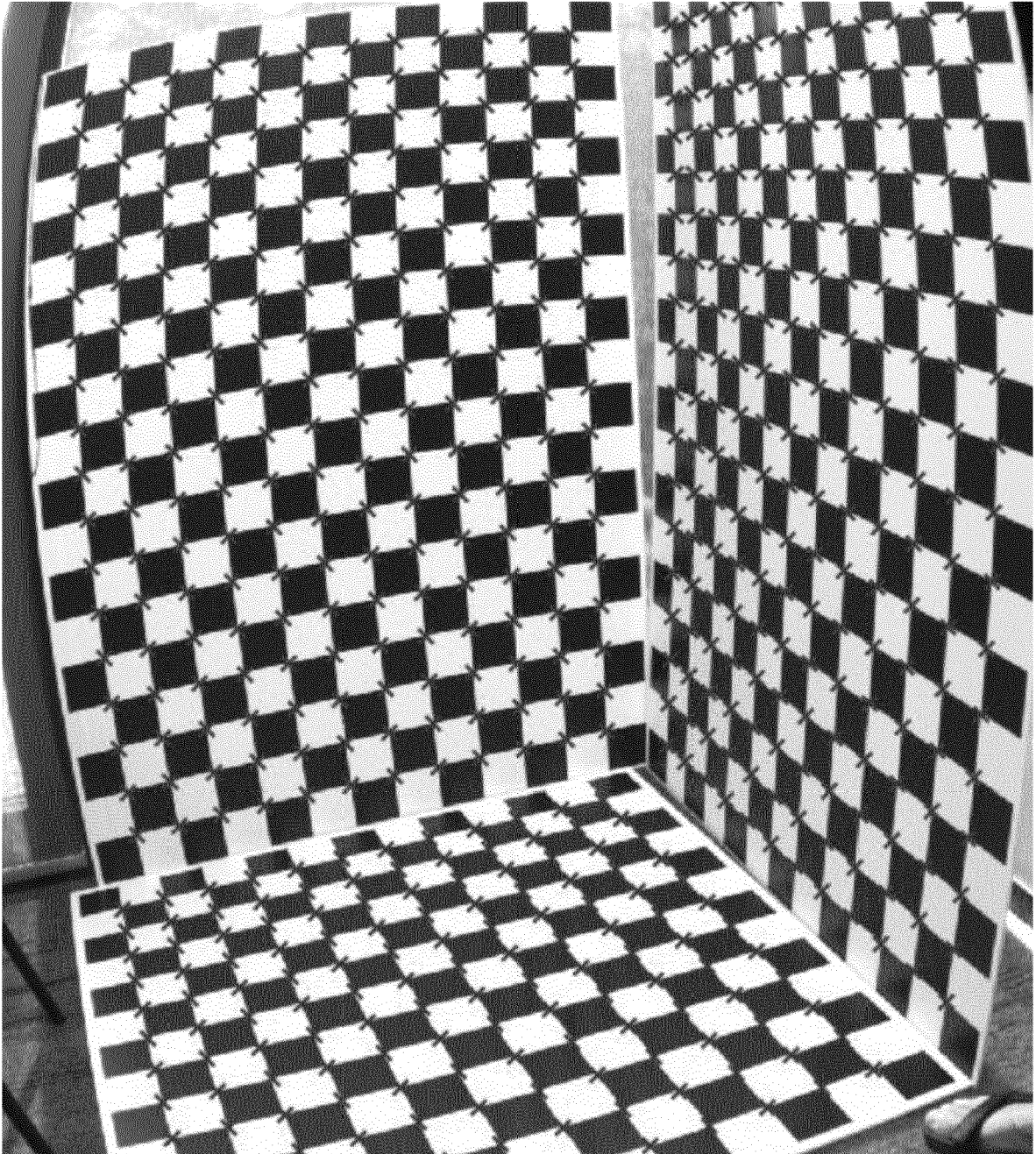


Fig. 1B

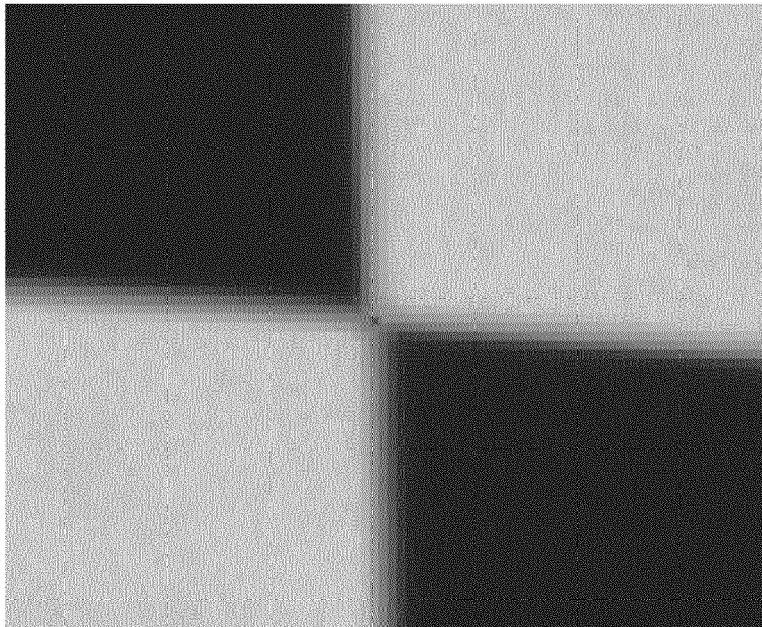
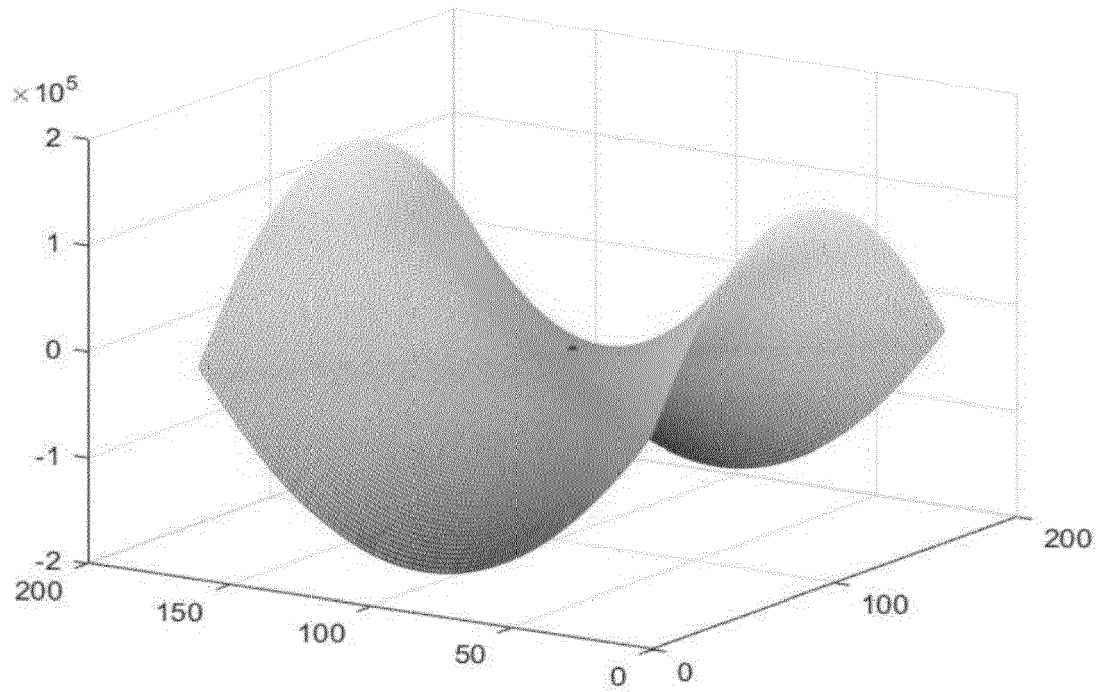


Fig. 2

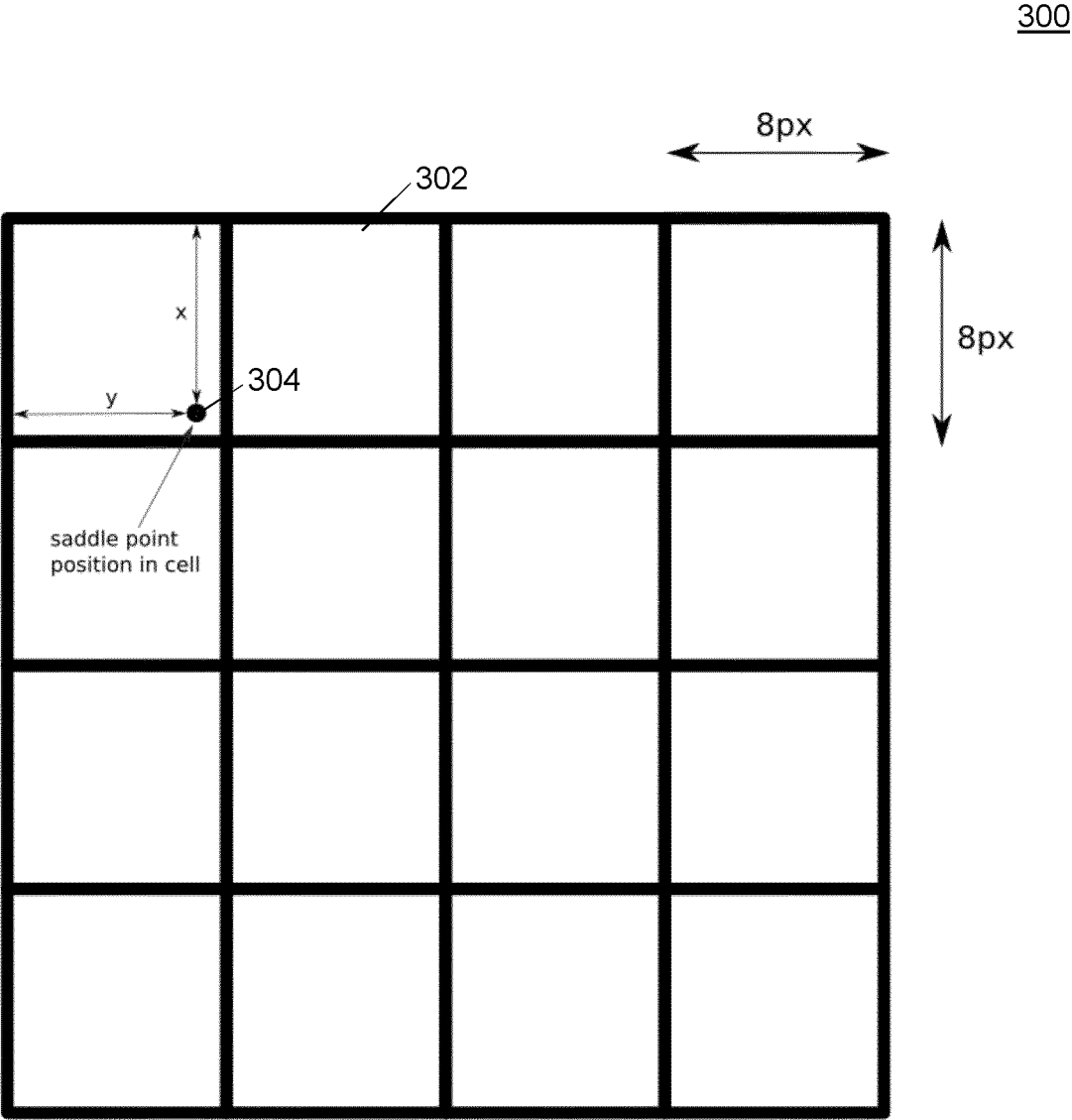


Fig. 3

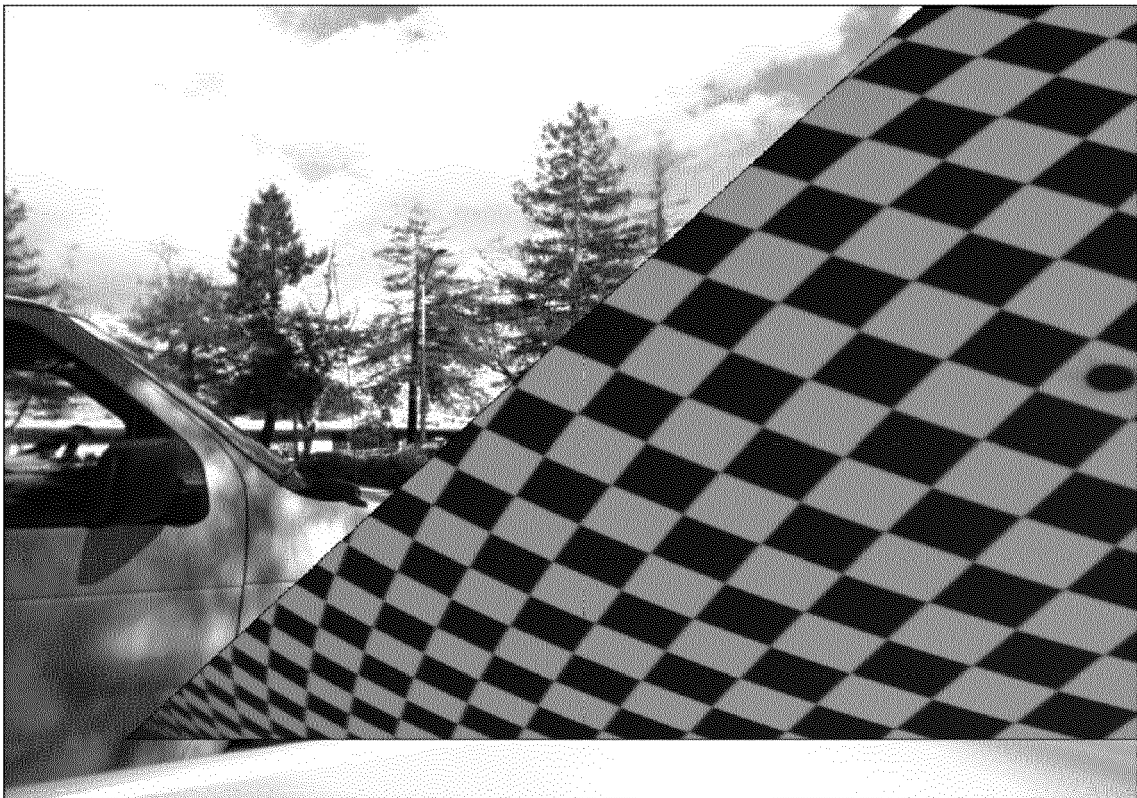


Fig. 4

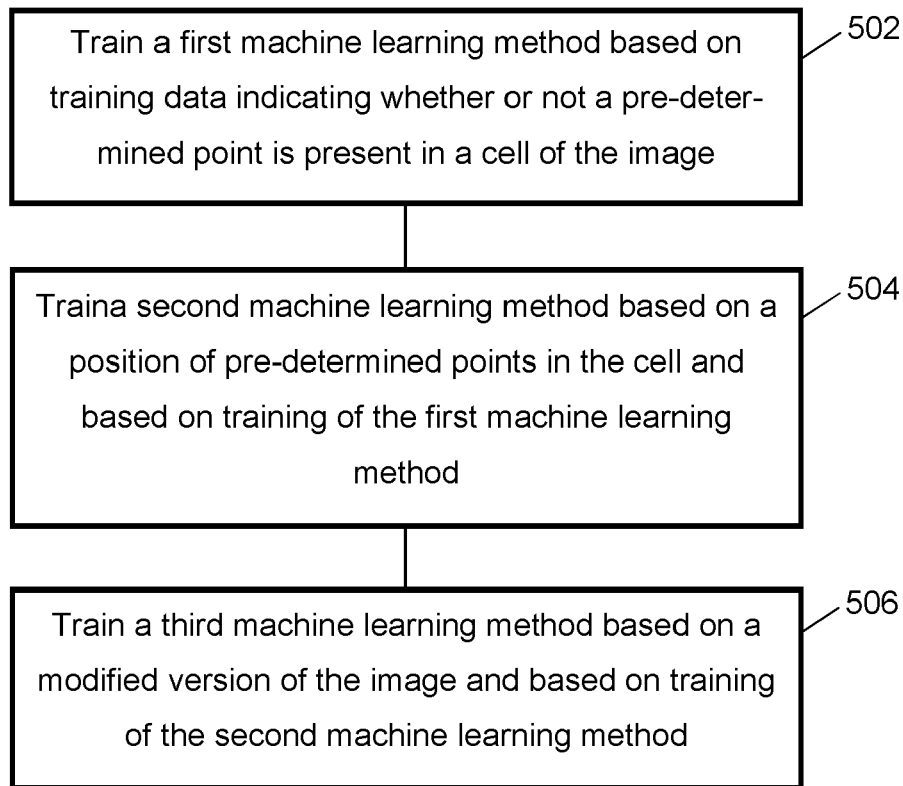
500

Fig. 5

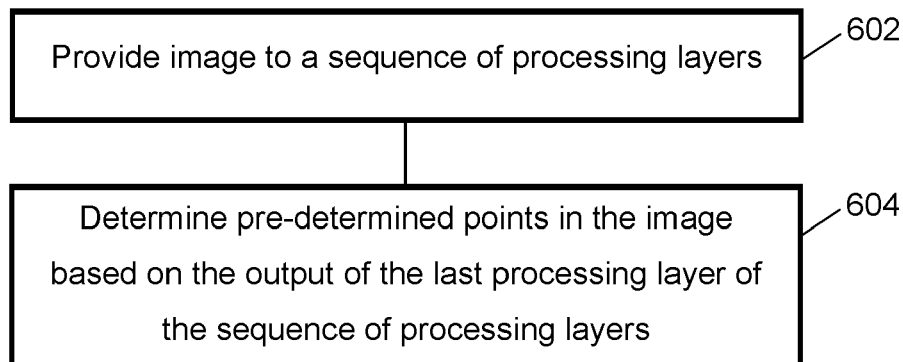
600

Fig. 6



EUROPEAN SEARCH REPORT

Application Number
EP 21 16 8669

5

10

15

20

25

30

35

40

45

50

55

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	DONNÉ SIMON ET AL: "MATE: Machine Learning for Adaptive Calibration Template Detection", SENSORS, vol. 16, no. 11, 1 November 2016 (2016-11-01), pages 1-16, XP055842716, Retrieved from the Internet: URL:https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5134517/pdf/sensors-16-01858.pdf>	8-15	INV. G06K9/62 G06N3/08 G06N3/04
A	* section 2 * * figure 3 * * figure 13 *	1-7	
A	----- RAZA SYED NAVID ET AL: "Artificial Intelligence based Camera Calibration", 2019 15TH INTERNATIONAL WIRELESS COMMUNICATIONS & MOBILE COMPUTING CONFERENCE (IWCMC), 24 June 2019 (2019-06-24), pages 1564-1569, XP055842870, DOI: 10.1109/IWCMC.2019.8766666 ISBN: 978-1-5386-7747-6 Retrieved from the Internet: URL:https://ieeexplore.ieee.org/stampPDF/getPDF.jsp?tp=&arnumber=8766666&ref=aHR0cHM6Ly9pZWVleHBsb3JlLm1lZWUub3JnL2RvY3VtZW50Lzg3NjY2NjY=> * sections II and III * * figure 2 *	1-15	
	----- -/--		
The present search report has been drawn up for all claims			
Place of search The Hague		Date of completion of the search 28 September 2021	Examiner Angelopoulou, Maria
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons ----- & : member of the same patent family, corresponding document	

EPO FORM 1503 03.82 (P04C01)



EUROPEAN SEARCH REPORT

Application Number
EP 21 16 8669

5

10

15

20

25

30

35

40

45

50

55

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	MARRERO FERNÁNDEZ PEDRO D. ET AL: "Fast and robust multiple ColorChecker detection using deep convolutional neural networks", IMAGE AND VISION COMPUTING, vol. 81, 1 January 2019 (2019-01-01), pages 15-24, XP055842882, GUILDFORD, GB ISSN: 0262-8856, DOI: 10.1016/j.imavis.2018.11.001 * section 2 * * figures 2-7,16,18 * -----	1-15	
			TECHNICAL FIELDS SEARCHED (IPC)
The present search report has been drawn up for all claims			
Place of search The Hague		Date of completion of the search 28 September 2021	Examiner Angelopoulou, Maria
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

 1
EPO FORM 1503 03.82 (P04C01)